THE INFLUENCE OF NURTURE UPON NATIVE DIFFERENCES

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN & CO., LIMITED
LONDON · BOMBAY · CALCUTTA
MELBOURNE

THE MACMILLAN CO. OF CANADA, LTD.
TORONTO

THE INFLUENCE OF NURTURE UPON NATIVE DIFFERENCES

BY

TRUMAN LEE KELLEY, PH.D.
PROFESSOR OF EDUCATION AND PSYCHOLOGY
IN STANFORD UNIVERSITY

New York
THE MACMILLAN COMPANY
1926

PRINTED IN THE UNITED STATES OF AMERICA

Set up and electrotyped. Published March, 1926.

Press of
J. J. Little & Ives Company
New York

PREFACE

Some years ago I felt that it would be a rather simple matter to determine experimentally the relative influences of heredity and environment upon mental performance. This appeared to involve merely a collection of ample data of a sort easily obtainable, and the analysis of such data by prosaic methods. As a beginning it seemed well to examine tentatively all the statistical steps which would arise in the undertaking. The attempt to do so proved disheartening because of the number of difficulties which it revealed. These were both logical and mathematical. To meet the mathematical difficulties certain new measures have been derived. They have been described in full in the December, 1925, issue of the *Journal of the American Statistical Association* and are used here for the first time in an experimental study. To meet the logical difficulties a number of explicit functional definitions are herein given.

Since the final issues of the study deal with heredity and environment it has been necessary to arrive at certain quantitative measures of these two factors, and at this point a great shortcoming (?) of the statistical method revealed itself. Whereas I had for years engaged in vigorous argument with colleagues as to the parts played in mental life by nature and nurture, I found (I might as well admit it now) that I had never to myself clearly defined either term and, still more surprising, that I could not find in the literature any precise functional definitions. The demand that these and other related concepts be precisely defined in behavioristic terms is a cruel infringement upon freedom of debate and for the moment seems to threaten

to take all the joy out of a problem which has long been an Elysian Field for the tired mind in a dogmatic mood.

Seriously, I expect that the first attitude of the reader of this study will be one of uneasiness that such rigid restrictions are placed upon the meaning of "nature," "nurture," "maturity," and "unit of mental measurement." I do not know that my definitions are ultimate, but I do maintain that they are serviceable and that only when these concepts are made very precise and related to human expressions of mental activity is their quantitative handling possible, and only then can the experimental method replace the purely dialectic one. The facts found and herein reported, and their social implications, are sufficiently momentous to point the value of an experimental study of heredity and environment.

I am indebted to the Commonwealth Fund for a grant which made the present investigation possible, and I also wish to express my obligation to Lewis M. Terman for a careful reading of the manuscript and for a number of constructive criticisms.

T. L. K.

STANFORD UNIVERSITY,
March, 1926.

CONTENTS

	PAGE
PART I: A STUDY OF AVERAGE CHILDREN . .	1
Section 1: FUNCTIONAL DEFINITIONS OF THE FACTORS ENTERING INTO AN ACHIEVEMENT SCORE . . .	1
Section 2: THE EXPERIMENTAL MEANS OF DIFFERENTIATING NATURE AND NURTURE FACTORS . . .	4
Section 3: THE BASIC UNIT OF MENTAL MEASUREMENT	8
Section 4: MEASURES OF IDIOSYNCRASY WHICH TAKE INTO ACCOUNT UNEQUAL RELIABILITY OF MEASURES	12
Section 5: THE ABSTRACTION OF THE NURTURE FACTOR	15
Section 6: FINAL MEASURES OF IDIOSYNCRASY AND OF NATURE AND NURTURE CONTRIBUTIONS TO IT . .	20
Section 7: THE EDUCATIONAL IMPORT OF CERTAIN CONDITIONS FOUND	23
Section 8: A REAFFIRMATION OF CERTAIN CRITICAL POINTS	29
PART II: THE IDIOSYNCRASIES OF GIFTED CHILDREN	32
PART III: APPENDIXES	38
Appendix A: VARIABILITIES OF GROUPS IN SENSED DIFFERENCES AND THE RELATION BETWEEN THE UNITS USED	38
Appendix B: BASIC DATA FOR NORMAL EIGHT, ELEVEN, AND FOURTEEN-YEAR-OLDS	46

THE INFLUENCE OF NURTURE UPON NATIVE DIFFERENCES

Part I. A STUDY OF AVERAGE CHILDREN

Section 1.—Functional Definitions of the Factors Entering into an Achievement Score

It is endeavored in this study to present a technique for the investigation of the relative influence of heredity and environment, and following this technique to report certain detailed findings as to the parts played by these two influences. The conclusions reached have been drawn from a study of achievement in elementary school subjects as measured by the Stanford Achievement Tests. As functional meanings of the terms heredity and environment become clear only as one studies their practical bearings, let us consider these concepts from the point of view of the adviser.

To be a wise counselor of youth, one should know the general level of abilities and the differences in the various abilities of the one advised; and further, it would be very valuable to know which of these differences are rooted in nature as a matter of heredity and which have been acquired. An instinctive or innate mental difference is of the nature of a bonus to its possessor and is a real and usable asset and foundation upon which he may build his future. Acquired differences may be equally or even more influential in determining one's future welfare, but that does not lessen the wisdom of utilizing such

unearned increments of human nature as are God-given to start with.

Presumably, any sort of acquired difference in abilities may be developed by training; thus by a special tutelage a child by original nature quite average in all respects might become, as judged by his peers, much more capable in botany than in zoology. The number and kind of such acquired specific differences is limited only by the peculiarities of the impressing nurture. Their variety is, probably, under the conditions of life today, very great. Inherited differences within the individual may be both less numerous and more amenable to classification than this quite limitless number of acquired differences. From much and varied evidence it seems to the writer to be so.

In order to isolate from other influences a single factor contributing to achievement, it is necessary that it be possible to state some unique manner in which this factor behaves or in which it influences an achievement score. We must thus set down all the influences which determine a score and define each in such a functional manner that it can be differentiated from the complex. A person's score in a test at any given time may be described as a product of his native ability, his maturity, his nurture, chance, and the units of measurement. The chance factor in a score has the unique characteristic of being uncorrelated with any of the other factors. This is the property which enables us to sift out its influence from the other influences. Maturity, for the purposes of this study, is a trait which is perfectly correlated, though not necessarily in a linear manner, with chronological age. As age varies the maturity values of each of two functions (x and y), being perfectly correlated with age, are perfectly correlated with each other. Thus for an age-heterogeneous population r (maturity in x) (maturity in y) $= 1$ while at the same time r_{xy} may have any value, depending upon the nature of the functions, between zero and one. The point here to be noted is that if we wish to eliminate maturity from consideration, then from our very definition of maturity all that

is required is that we deal with deviations from age means. We will designate such measures:

by $x_1, x_2, \ldots$ for an 8.00–8.99-year-old group,
by $\mathbf{x}_1, \mathbf{x}_2, \ldots$ for an 11.00–11.99-year-old group,
by $X_1, X_2, \ldots$ for a 14.00–14.99-year-old group.

Throughout we will follow the same scheme in designating other measures; italic type will refer to 8-year-olds, bold face type to 11-year-olds, and capitals to 14-year-olds.

Should the reader object to the definition of maturity here employed, he may substitute some other word for the thing defined. Given groups of American white children at the average school grade for their age, it is a well-known fact that there is an actual increase in mean score from year to year. It is this that is called maturity because it has the outstanding characteristic of agreement with chronological age, and it is the effect of this which is eliminated by the procedure adopted. The word maturity in this connection seems appropriate to the writer, and he consistently uses it with this and no other meaning. Should one dislike the term, let him coin or choose a word to his liking and substitute it throughout with no loss whatever to the argument.

By the same definition of maturity, any common heredity, such for example as that due to all the subjects being American white children; and any common nurture, such for example as that due to their having attended American public schools, are immediately ruled outside of the definitions of heredity and nurture. These common causes are incorporated in the definition of maturity and are made constant by the simple device of dealing with deviation scores. Insofar as schooling differs and has a different effect upon children of the same age and ability and in the same grade, the effect is nurture. In other words, it is only insofar as the effect of this common schooling is the same that it is incorporated into maturity.

Nature and nurture are then defined as influences which, under

the conditions of American white school children in the public schools, all in the same school grade and of the same chronological age, still cause differences in achievement to be manifest. Someone may claim that nurture is too narrowly defined, — that we should compare 14-year-old children who have attended public schools with 14-year-olds brought up in the jungle to determine the influence of nurture; another may claim that we have defined nature too narrowly, — that we should compare the achievement of a cat in a puzzle-chamber and a human being in the same chamber to determine the influence of nature. To both it may be answered that this study is of influences at work under conditions typical of American school life and of American white children. The guidance counselor is not ordinarily concerned with such broad differences in either nurture or nature as those first mentioned. The native of the jungle is not a rival of the schoolman; nor the cat of *homo sapiens;* nor even, indeed, the young child of the old one. Within the narrow limits defined — American white children of a given age trained in the public schools — there exist native differences and nurture differences which show themselves in substantial achievement differences and keen rivalries, and an answer to the problem of heredity and environment within these limits is the issue of moment, for it is the "felt" issue of the child himself.

Accordingly, nature is defined as a trait, making for individual differences within this field, which does not change with age, and nurture is defined as an influence, uncorrelated with nature and making for individual differences within this field, which changes with the length of time or number of years through which it acts.

Section 2.—The Experimental Means of Segregating Nature and Nurture Factors

At any given age the child's true score as a deviation from his group mean may be set equal to $a + u$ (nature + nurture), and

at a subsequent age equal to $a + qu$ (the same nature factor + a changed nurture factor).

If we designate a person's true score as a deviation from his group mean at age $8\frac{1}{2}$ as s and his deviation score at age $11\frac{1}{2}$ as $\mathbf{s}$, we then have:

$$s = a + u \qquad [1]$$

$$\mathbf{s} = a + qu \qquad [2]$$

Thus, if we know q, and if we have the scores s and $\mathbf{s}$ in appropriate units, we can solve these two equations for a and u, the nature and nurture factors contributing to the individual's score at age 8.

This very simple statement gives the line of argument of this treatment, but the actual setting down of two such equations as [1] and [2] cannot be done until a number of difficulties have first been surmounted: (a) the multiplier of the nurture factor, q, must be found; (b) we do not have true scores, and the achievement scores which are available all have substantial chance factors, so that the procedure adopted must eliminate any systematic error from this source; (c) it should not be assumed that the units of any existing tests are appropriate as they stand; (d) it has not in this experiment been possible to test the same children at two ages, in which case the variability or scatter of the nature factor for the 8-year-olds (represented by σ^2_a, — the standard deviation squared, or the variance* of the nature factor of the individuals composing the group) would be identically the same as at the later age. It has therefore been necessary to exercise caution in selecting age groups in order to insure a known relationship between the variabilities at the different ages.

In considering this last point (d) it is to be noted that it is very difficult to select a truly random age group from a school system except possibly for the ages 11, 12, and 13. A random

* Following R. A. Fisher ("The Correlation Between Relatives on the Supposition of Mendelian Inheritance," *Transactions of the Royal Society*, Edinburgh, 1918) the term variance is used to mean σ^2, the standard deviation squared.

sampling of 8-year or 14-year-olds cannot be obtained by selection from the elementary school, as many 8-year-olds have not entered school and many 14-year-olds have left it. For many children of these ages the elementary school environment is lacking and the conditions of our problem not met, as we have set out to study conditions of children when compared with their peers. Accordingly a complete age sampling has not been attempted, and in place thereof, normal or typical age groups have been used.

By a study of records of approximately 10,000 elementary school children throughout the United States, it was found that the average school grade for the 8-year-olds was 2.79*; for the 11-year-olds 5.65; and for the 14-year-olds 8.53. Groups of children located in school within one-half of a grade and within one-half of a year in age were selected, with a view to meeting experimentally these grade and age locations. The groups actually selected are as follows:

	8-year-olds	*11-year-olds*	*14-year-olds*
Population	825	887	982
Ages between	8.00–8.99	11.00–11.99	14.00–14.99
Mean school grade	2.80	5.65	8.55
σ of school grade	.275	.270	.260
Mean St. Ach. total score	109	384	569
σ of total scores	57.75†	76.38†	73.04†

* 2.0 means the beginning of the second grade; 2.5 the middle of the second grade, etc.

† The Stanford Achievement Test has six parts—Paragraph Meaning, Sentence Meaning, Word Meaning, Arithmetic Computation, Arithmetic Reasoning, and Spelling—for grades 2 and 3, and nine parts—the same six plus History and Literature Information, Science Information, and Language Usage—for grades above the third. In this treatment the total score, no matter which age group is considered, means the total score on the six parts only. The standard deviations reported were computed by the formula

$$\sigma = \frac{D}{2.5631}$$

in which σ is the standard deviation desired of the total scores for the group in question and D is the 10–90 percentile range. The relationship holds for normal distributions. The standard deviation, σ, was computed via D to shorten the labor. Any slight error in thus determining σ is inconsequential in comparison with other errors which enter into this problem at other points.

The children of each group vary widely one from another in achievement in reading, arithmetic, spelling, etc., but are very homogeneous in the matter of grade and age. Individual differences within an age group so selected are to be attributed to nature and to type of nurture rather than to differences in the number of years through which nurture has acted. This is as we wish it.

If 8-year-olds in the second grade are followed through the school, they will not all be found in the fifth grade at 11, nor in the eighth grade at 14. The 8-year group is presumably natively more heterogeneous than the other groups, so that even if our units are appropriate, σ_a for the 11-year-olds is not identical with σ_a for the 14-year-olds and we may not write down as true:

$$\sigma^2 = \sigma_a^2 + \sigma_u^2$$
$$\boldsymbol{\sigma}^2 = \sigma_a^2 + q^2\sigma_u^2$$

Having so selected the groups that σ_a^2 for the one age does not equal σ_a^2 for the other, we must discard the use of the variabilities in x_1 and $\mathbf{x}_1$ measures as providing the basis of comparison.

If, instead of dealing with abilities of pupils, we use differences of abilities within the individuals we thereby get away from the specific trait upon the basis of which selection has in part been made. Thus if for the entire 8-year group σ_δ^2 is a measure of the tendency to be different in Trait (1) from Trait (2), and if this difference is expressed as a function of nature and nurture thus:

$$\sigma_\delta^2 = \sigma_a^2 + \sigma_u^2 \qquad [3]$$

we may, if our units are appropriate, write for the 11-year group:

$$\boldsymbol{\sigma}_\delta^2 = \sigma_a^2 + q^2\sigma_u^2 \qquad [4]$$

The statement herein that σ_a^2 of the first equation is equal to σ_a^2 of the second equation is stating that having typical children in each instance the native tendency toward idiosyncrasy is the same. The specific measure used, δ, is defined later.

Section 3.—The Basic Unit of Mental Measurement

We must now consider what constitutes appropriate units of measurement. The geneticist might advocate such units as are proportional to genetic differences.

In Mendelian studies of the dimension of some organ, say a limb, it is supposed that a number of genes constitute the underlying cause of the measure. It is assumed that if all nurture factors — food, sunlight, etc.— are constant throughout, then the genetic factors completely determine the measured trait. If three equally potent genes are involved and if each can be influential in one of two degrees: A, a; B, b; C, c; we may have the following genetic influences affecting the length of the limb:

A, B, C		giving the greatest length of limb
A, B, c A, b, C a, B, C	indistinguishable one from another in effect	giving a shorter limb
A, b, c a, B, c a, b, C	indistinguishable one from another in effect	giving a still shorter limb
a, b, c		giving the shortest length of limb

If these four limb lengths, beginning with the shortest, have numerical values t, $(t + k)$, $(t + 2k)$, and $(t + 3k)$, then a measured somatic difference of k would correspond to a unit genetic difference. Only when such a statement of obtained limb lengths is possible does the biologist have for a graduated trait genetically significant units of measurement. As biologists would not ordinarily suppose that each of three genes were equally potent in determining a quantitative trait such as limb length, the present statement may seem crude, but the point here made could equally well, though not so briefly, be made starting with the assumption of different quantitative influences attaching to each of the three genes.

So far as the writer is aware, length expressed in inches or

fractions thereof has never been found to be expressed in genetic units, nor has some definite function of inches been found to correspond to a genetic unit. It is thus seen that, from the genetic point of view, neither the inch nor any as yet defined function thereof has claim to genetic significance. A similar situation holds with reference to our units of weight and time. The writer claims that the sense difference is *per se* the only meaningful unit of measurement of mental phenomena, and he has presented the present argument merely to show to the reader that if "natural" units are defined in genetic terms, then the sense-difference is in a neither better nor worse position than are such long-established measures as the inch, pound, and second, for in each case the relation of the used unit to the genetic unit is a thing to be ascertained.

If competent judges appraise Individual A as being as much better than Individual B as Individual B is better than Individual C, then it is so, as there is no higher authority to appeal to. If sensed difference units have no relationship with genetic differences, we would then indeed be in hard luck, for this would be equivalent to the declaration that mental genetic influences carry with them no correlates capable of being sensed. Such a conclusion is hardly conceivable in view of the known genetic influences upon stature, eye color, etc., etc., all when measured in units approximately equal to or readily related to sensed units.

The following paragraphs quoted from Kelley (1923)*, page 418, deal with the question here considered. "It might seem axiomatic that there cannot be a science of quantitative measurement until and unless there is established a particular unit of measurement. This is, however, true only in a limited sense; for it is quite conceivable that one could have a science of physical phenomena in which the units were such that the scale of time intervals was the square of the present intervals measured

* Kelley, T. L., "The Principles and Technique of Mental Measurement," *American Journal of Psychology*, July, 1923.

in seconds, and in which the length scale was logarithmic as compared with the present scale in centimeters, etc. Of course, in terms of these new units, all the laws of physics would be stated by means of formulas different from and in general more cumbersome than our present formulas; but nevertheless we could have an exact science. The existence of the science does not lie in the units employed, but in the relationships which are established as following after the choice of the units.

"A parallel situation holds with reference to mental measurement; so that, starting with units however defined, if we can establish important relationships between phenomena measured in these units, we have proceeded scientifically. The choice of the unit is purely a question of utility. It is preferable so to define it as to lead to the simplest possible algebraic statements of the important relationships." As the initial and the end terms in the thinking process dealing with mental life are phenomena as sensed, we immediately eliminate two steps in all algebraic treatment when we employ in the first instance and throughout to the conclusion units which are sensed as equal.

Even in a field where other units are firmly established we resort to sense difference units when a novel situation arises. Imagine a traveller in a foreign land, viewing pygmies and giraffes for the first time, and asking himself which are the more variable. He would spontaneously think in terms of proportion. The ratio of giraffe G to giraffe J is sensed and compared with the ratio of pygmy M to pygmy N. He would not think in terms of inches unless established habits made such a procedure serviceable in determining ratio differences. Weber's law that sense differences are proportional to the logarithms of physical differences is an expression of the fact that physical differences are not appreciated except as they become sense differences. Dealing with the small range in height commonly found among civilized adults, it is for most thinking purposes immaterial whether logarithms or actual measures are considered, but in dealing with differences as great as those between

pygmies and giraffes, the inch is secondary to the sense difference.

The sense difference may be employed wherever there is a sensed sameness in function. Thus, if competent judges affirm that reading ability as found in young and old, dull and talented, is comprehensible as a single trait, and if 75 out of 100 competent teachers state that 8-year-old A is better in reading than 8-year-old B, and if the same or 100 other equally competent teachers compare two 14-year-old children, X and Y, with the result that 75 conclude that X is better in reading than Y, then X is just as much better in reading than Y as A is better than B. Though A and B reveal their abilities upon "I see the cat. It is a big cat," etc. and X and Y reveal their abilities by reading Shakespeare, there is no appeal (except by using larger numbers of competent judges) from the decision that in reading A is just as superior to B as X is to Y.

In other fields (height, weight, temperature, time, etc.) we have related sensed differences to extra-sense standards and units, found the relations between them relatively simple, and thus the standard, extra-sense in the first instance, becomes comprehensible and in fact a sense standard. As a consequence of the procedure in non-mental fields there are probably many who feel the inclination to do the same thing in dealing with mental phenomena. However, in this field these extra-sense standards are not established. Obviously, if established, they would only become meaningful as they were related to sense standards. We are then attacking the problem most directly by beginning with the sense difference and keeping it throughout as the basic unit of measurement, since it is the thing that is thought of when mental differences are discussed.

We find (see Appendix A) that in estimated true sense differences the standard deviations of total scores of the 8, 11, and 14-year-old groups are in the ratio .982 : .895 : 837. Let us assume, — feeling sure that the assumption is only a first approximation and one that should be investigated in a more

detailed study, — that the ratio between total achievement abilities of the three groups is the same as the ratio between abilities in the separate subjects: Paragraph Meaning, Sentence Meaning, Word Meaning, Computation, etc. Under these conditions if the score of a certain 8-year-old in Paragraph Meaning, expressed as a deviation from his age mean, is x_1 and if the standard deviation of 8-year-old Paragraph Meaning scores is σ_1, then $(.982)\ x_1 / \sigma_1$ is a deviation in units proportional to sense differences. If similarly x_2 and σ_2 refer to Word Meaning, then $(.982)\ x_2/\sigma_2$ is the deviation in comparable units and

$$.982\left(\frac{x_1}{\sigma_1} - \frac{x_2}{\sigma_2}\right)$$

which we will call $.982\ d_{12}$, is a measure of individual idiosyncrasy in sense difference units.

Section 4.—Measures of Idiosyncrasy which Take into Account Unequal Reliability of Scores

The preceding argument holds provided x_1 and x_2 are equally reliable. Since, however, the various Stanford Achievement tests are not equally reliable, a modification of this treatment is necessary. We can derive a measure of idiosyncrasy of this same general sort yielding results as though the tests were equally reliable and equal to some assigned value. The value which will be assigned is .80 for the 14-year-old group and comparable values* as given by the equation for other groups as follows:

$$\frac{\sigma_\infty}{\boldsymbol{\sigma}_\infty} = \sqrt{\frac{r_{1I}}{1 - r_{1I}}} \Big/ \sqrt{\frac{\mathbf{r}_{1I}}{1 - \mathbf{r}_{1I}}} \qquad [5]$$

* If the true variabilities of the 8 and 14-year-old groups are in the ratio .982: .837 and if the reliability of a single test, such as the Paragraph Meaning test, when given to the 14-year-old group is .80 (which is practically the average of the obtained reliabilities of all the tests for this group), then the reliability for the 8-year group is given by the equation

$$\frac{.982}{.837} = \sqrt{\frac{r_{1I}}{1-r_{1I}}} \Big/ \sqrt{\frac{.8}{1-.8}}$$

(See Kelley, *Statistical Method,* Formula [177].) Solving we obtain $r_{1I} = .85$. We find similarly for the 11-year-olds $r_{1I} = .82$.

.82 for the 11-year-old group and .85 for the 8-year-old group. Instead of dividing each deviation measure by the standard deviation of the group let us divide by the estimated true standard deviation of the group.* Let σ_∞ equal the estimated true standard deviation of the 8-year-olds in Paragraph Meaning and σ_ω the same in Sentence Meaning. We have as a measure of individual idiosyncrasy $x_1/\sigma_\infty - x_2/\sigma_\omega$, which we will designate by the symbol Δ_{12}. Let us express each difference as a multiple of its standard error, thus, $\Delta_{12}/\sigma_{\Delta_{12}\cdot\infty\omega}$. For simplicity let us call this quotient δ_{12}. The variance of δ_{12} is a measure of the extent to which the members of the entire group tend to be different in Paragraph Meaning and Sentence Meaning. More exactly expressed, the excess of $\sigma^2_{\delta_{12}}$ over 1 is such a measure, for if the two traits are identical so that there is in reality no idiosyncrasy whatever, then $\sigma^2_{\delta_{12}}$ equals 1. We will accordingly call $(\sigma^2_{\delta_{12}} - 1)$ the group measure of idiosyncrasy and designate it by the symbol i^2_{12}. Thus we have†

$$i^2_{12} + 1 = \sigma^2_{\delta_{12}} = \frac{\dfrac{1}{r_{1\text{I}}} + \dfrac{1}{r_{2\text{II}}} - 2r_{\infty\omega}}{\dfrac{1}{r_{1\text{I}}} + \dfrac{1}{r_{2\text{II}}} - 2}. \qquad [6]$$

Since the correlation between true Paragraph Meaning scores and true Sentence Meaning scores does not depend upon the reliability of instruments of measurement, our measure of it, $r_{\infty\omega}$, will not change in any systematic manner no matter what the values of $r_{1\text{I}}$ and $r_{2\text{II}}$. This fact may be utilized in comparing measures of idiosyncrasy obtained from different populations. The group idiosyncrasy for the actual tests used is as given by this last formula wherein $r_{1\text{I}}$ and $r_{2\text{II}}$ are the actual reliabilities of the tests in question, but we may immediately find what the group idiosyncrasy would be for tests of different reliabilities by substituting the required values for $r_{1\text{I}}$ and $r_{2\text{II}}$. If we desire

* See T. L. Kelley, "Measures of Correlation Determined from Groups of Varying Homogeneity," *Jour. Am. Statis. Assn.*, Dec., 1925.

† Ibid., Formula [15].

a measure of group idiosyncrasy for tests of equal reliability we have $r_{11} = r_{2\mathrm{II}}$ and Formula [6] becomes

$$i^2_{12} = \frac{r_{11}}{1 - r_{11}} (1 - r_{\infty\omega}). \qquad [7]$$

Obviously in case $r_{\infty\omega}$ does not equal 1, the size of this measure, i^2_{12}, depends upon the reliabilities of the tests used. This simply means that if there is any difference between two mental functions it is poorly established if the tests have low reliability, and very well determined if the instruments of measurement are highly accurate. In comparing the idiosyncrasies for different groups and between different tests from a single group, it is essential that comparable reliabilities be employed. In a preceding paragraph it was shown that these reliabilities are .85, .82 and .80 for the 8, 11 and 14-year groups respectively. Substituting these values for r_{11} in Formula [7] we have as comparable measures of group idiosyncrasy:

$i^2_{12} = 5\frac{2}{3}(1 - r_{\infty\omega})$ [Measure of idiosyncrasy of typical 8-year-olds] [8]

$\mathbf{i}^2_{12} = 4\frac{5}{9}(1 - \mathbf{r}_{\infty\omega})$ [Measure of idiosyncrasy of typical 11-year-olds] [9]

$\mathrm{I}^2_{12} = 4(1 - \mathrm{R}_{\infty\omega})$ [Measure of idiosyncrasy of typical 14-year-olds] [10]

The measures i^2_{12}, $\mathbf{i}^2_{12}$, and I^2_{12} may be immediately compared with each other for the ratio $i_{12}/\mathbf{i}_{12}$ is independent of the particular reliabilities of tests used when these reliabilities are in the ratio given by Formula [5]. This ratio may be expressed as follows:

$$\frac{i^2_{12}}{\mathbf{i}^2_{12}} = \frac{\left(\dfrac{r_{11}}{1 - r_{11}}\right)(1 - r_{\infty\omega})}{\left(\dfrac{\mathbf{r}_{11}}{1 - \mathbf{r}_{11}}\right)(1 - \mathbf{r}_{\infty\omega})} = \frac{.982\ (1 - r_{\infty\omega})}{.837\ (1 - \mathbf{r}_{\infty\omega})}.$$

Accordingly, if r_{11} and $\mathbf{r}_{11}$ are always chosen, as they should be, in the proportion given by Formula [5], then we finally arrive at measures of idiosyncrasy i^2_{12}, $\mathbf{i}^2_{12}$, and I^2_{12}, which maintain pro-

portions to each other entirely independent of the reliabilities of our instruments of measurement. To secure this freedom from the particular instruments used was the reason for using Δ_{12} instead of d_{12} measures.

To summarize the effects of the steps in procedure thus far determined: (a) Dealing with deviations from age means eliminates maturity. (b) Dealing with sense differences (which yielded the ratio .982 : .895 : .837) eliminates the particular units of measurement and substitutes sense difference units. (c) Dealing with i^2_{12} measures of group idiosyncrasy (as related to other i^2_{12} measures) eliminates any systematic effects of chance. (d) Dealing with large typical populations eliminates much of the chance effects of chance. We thus have left in these i^2_{12} measures of idiosyncrasy the effects produced by differences in nature and nurture plus a small chance factor which it is impossible to eliminate.

To distinguish between nature and nurture it is necessary to utilize a behavior property in which they differ. As stated earlier this property is that the effect of nature remains constant with a change in age while the effect of nurture changes with age according to a certain law. It is hereby attempted to give an approximate statement of the relation between nurture and age.

Section 5.—The Abstraction of the Nurture Factor

The reasoning followed is based upon the common argument that the longer nurture acts, the greater is its effect. Eight, eleven and fourteen-year-olds have been studied, it being held that if u is the total life nurture effect* some fraction of this, qu, is the effect of nurture in the case of the 8-year-olds, some greater fraction, $\mathrm{q}u$, will be the effect for the 11-year-olds, and

* Earlier in this article u was given as the 8-year-old nurture effect. From here on it represents the total life nurture effect, that is, the sum of all the nurture influences up to adulthood, at which time it is assumed that stability in mental functions has been reached.

some still greater multiple, Qu, will be the effect for the 14-year-olds. The values of q, $\mathbf{q}$, and Q finally arrived at are: q = .262, $\mathbf{q}$ = .679, and Q = .922. The considerations leading to these results are set forth herewith.

It can scarcely be maintained that nurture in its influence upon such capacities as reading, arithmetic, and spelling has had equal opportunity per year to influence an average child from age zero up to age 14 or beyond, and that the nurture influence up to age 14 is $\frac{14}{8}$ that up to age 8. The value of this fraction will undoubtedly be greater than $\frac{14}{8}$. It seems much more reasonable to maintain that the nurture influence per year is proportional to the rate of change of the function involved. Thus a mental function which has not been developing or growing at a given age, for example spelling at age 1, is not at that age being affected by nurture, while a function rapidly growing is amenable to much influence by nurture. If we consider this influence as proportional to the rate of growth of the function, then the size of the function, measured from a true zero point at any given age, is proportional to the sum of all the preceding nurture influences up to that age.

Even if we consider this problem with reference to an extreme case we find it not at all unreasonable. Let us choose height, a trait which is probably mainly a matter of nature and maturity. Height is rapidly changing at ages before ten and at puberty. If the nurture influence is proportional to the rate of change of the function, it would be considered some two times as great at these ages as age 11 or 12. Now, this may very well be the case, for even if nurture in the matter of height is nearly negligible at age 11 or 12, it will remain insignificant when we multiply it by two to secure the nurture value for age 14. This is simply an illustration showing that the assumption that the nurture influence is proportional to the rate of change of the function in no wise assumes that in itself it is either a great or a small influence in comparison with nature and maturity.

Following the principle laid down we have for age 8½ an aver-

age Stanford Achievement score of 19.95 above zero (See Table E); considering the same tests, for age 11½ a score of 51.61; for age 14½ a score of 70.12; and for the average adult 76.02. Thus the nurture influences for the 8-year-olds would be 19.95/76.02; for the 11-year-olds 51.61/76.02; and for the 14-year-olds 70.12/76.02 of the average adult influence. These ratios provide the desired nurture factors:

$$q = .26,\ \mathbf{q} = .68, \text{ and } \mathrm{Q} = .92$$

Again, if we consider that the nurture influence, as affecting reading, arithmetic, etc., for the 8-year-olds is equivalent to two effective years; for the 11-year-olds equivalent to five effective years; and for the 14-year-olds to eight effective years, then the nurture influence for the 11-year-olds is 5/2 and for the 14-year-olds it is 8/2, that of the 8-year-olds. These ratios do not vary greatly from 51.61/19.95 and 70.12/19.95. Thus, both by the argument based upon size of the function and by that upon number of years of instruction, factors close to the value recorded are indicated.

We now have an expression of the relative importance of nurture in creating a score, or in creating an idiosyncrasy, for these three ages. These ratios provide the necessary data for building up the equations from which a and u, the nature and life nurture factors, may be obtained.

$$i^2_{12} = a + .26u \qquad [11]$$

$$\mathbf{i}^2_{12} = a + .68u \qquad [12]$$

$$\mathrm{I}^2_{12} = a + .92u \qquad [13]$$

The two magnitudes a and u are determined from these three equations by the method of least squares.

As an illustration let us consider Arithmetic Reasoning and Spelling represented by scores x_5 and x_9. Then $r_{\infty\omega} = .60$; $\mathbf{r}_{\infty\omega} = .50$; and $\mathrm{R}_{\infty\omega} = .42$ (from Tables F, G, and H): and $i^2_{59} = 2.29$; $\mathbf{i}^2_{59} = 2.28$; and $\mathrm{I}^2_{59} = 2.33$ (from Equations 8, 9, and 10). Solving by least squares for a and u yields $a = 2.27$

and $u = .06$. The absolute values of a and u depend upon the reliabilities of the instruments of measurement, but the proportion between them does not. We may therefore conclude that approximately 97 per cent of the adult difference between Arithmetic Reasoning and Spelling abilities is to be attributed to original nature. Thus nurture has in the long run negligibly influenced the native idiosyncrasies in abilities in these two fields.

Someone may ask what meaning there can be in the term "native" when applied to such abilities as Arithmetic Reasoning and Spelling, in view of the fact that the very young child is possessed of zero ability and therefore of zero difference in ability in each. This is a puzzle, but it is of the same sort as that of two infants, a boy and a girl, one of whom later develops a moustache and the other does not. In this case both in the beginning manifest zero achievement and zero difference in the matter of moustaches. As we do not hesitate to attribute such a developed difference to nature, so here we need not hesitate to attribute a mental difference appearing late in life to nature.

Another may be puzzled by the fact that the native idiosyncrasies are expressed in terms of sense differences when as a matter of fact the very young child shows neither trait and therefore no difference is sensed. If it be recalled, however, that the sense difference is divided by the standard error of such difference, we are of course only concerned with the quotient, sense difference divided by the error of judgment, and in judging infants both the numerator and denominator approach zero, so that the quotient may approach any value consistent with the facts of mental life when first discernible. There is no necessity that these quotients shall approach zero as earlier and earlier chronological ages are considered. In further support of this view we may note that child observation may easily suggest the existence in the case of certain children of a spelling ability (possibly indicated by the repetition of sounds heard) before the existence of an arithmetic reasoning ability (indicated by the

deduction that something is missing if only one of two rattles is in hand) and vice versa with other children. Thus, in the first manifestation the thing sensed by the parent may be a difference in abilities. It is rather unprofitable to speculate upon this matter, for the errors of judgment are undoubtedly large with respect to the traits judged at the early age in question, but the writer finds no logical or statistical argument requiring that children at birth or at any other age must be considered to possess zero idiosyncrasy with reference to the traits measured by the Stanford Achievement Tests.

Nurture has so long been thought of as something added to nature that the fact that in general the native differences between arithmetic reasoning and spelling abilities have been increased by nurture seems natural, but upon further thought we may ask, why should this be so? If the infant in actual and potential capacities is not a blank, and he certainly is not, then nurture has an opportunity to level differences as well as to create them, all depending upon the quality of the nurture. Much of public education is with a view to leveling differences, curbing anti-social mutations, developing a common culture, and making all men equal. We will in fact find that the data presented in this article point to the nurture effect as being more often a leveling of native differences than an augmenting of them. Whether this is "good" or "bad" is an open question, but certainly the data tend to support Dewey's doctrine that the young child has more individuality than the older child and that we are accomplishing by education what certain democratic educators earnestly desire — the creation of a homogeneous citizenry. Though we may not go as far in this direction as they would like, still such advocates should find satisfaction that on the whole the nurture influences which operate upon typical children seem to wash out individual peculiarities rather than augment them. This general conclusion will be further examined in Part II in connection with the nurture of gifted children.

Section 6.—Final Measures of Idiosyncrasy and of Nature and Nurture Contributions to it

The raw data for determining *a* and *u* measures are given in the appendix, Tables F, G, and H. From these, Table I, giving in each cell i^2_{12} values in order for 8, 11, and 14-year-olds has been obtained. Rigorously established probable errors of these values are not at present available. The writer estimates that the errors due to technique and due to sampling give a final probable error between .1 and .2.

TABLE I

Comparable Idiosyncrasy Measures for Typical 8, 11 and 14-Year-Olds

	Par. Mean.	*Sent. Mean.*	*Word Mean.*	*Computation*	*Arith. Reas.*	*Sci. Inf.*	*Hist. + Lit.*	*Lang. Usage*
Sent. Mean.	1.12 1.18 .85							
—Word Mean.	.78 1.17 .59	1.08 .68 .47						
Computation	2.77 2.50 2.39	3.00 2.60 2.38	2.99 2.79 2.62					
Arith. Reas.	1.85 1.72 1.69	1.78 1.89 1.83	2.03 2.09 1.90	2.28 1.66 1.00				
Sci. Inf.	1.33 1.34	1.13 1.37	.92 1.11	2.59 2.63	1.66 1.56			
Hist. + Lit.	1.40 1.32	1.13 1.25	1.07 .82	2.56 2.75	1.82 1.80	.94 .71		
Lang. Usage.	1.22 1.16	.93 .92	1.10 .86	2.37 2.38	1.80 1.99	1.32 1.60	1.25 1.39	
Spelling	1.33 2.02 1.59	1.80 1.70 1.63	1.62 1.53 1.45	2.80 2.58 2.24	2.29 2.28 2.33	1.95 2.40	2.04 2.29	1.58 1.39

From Table I the values in Table II, not involving Science Information, History and Literature Information, or Language Usage, were calculated by the method of least squares, from equations [11], [12], and [13]. The values involving these three tests were directly determined from equations [12] and [13], since equation [11] was not available.

TABLE II

NATURE AND NURTURE CONTRIBUTIONS TO ADULT IDIOSYNCRASY (Science Information, History and Literature Information, and Language Usage figured from 11 and 14-year-old data only)

	Par. Mean.	*Sent. Mean.*	*Word Mean.*	*Computation*	*Arith. Reas.*	*Sci. Inf.*	*Hist. + Lit.*	*Lang. Usage*
Sent. Mean.	1.26 −.34 .92							
Word Mean.	.94 −.15 .79	1.32 −.93 .39						
Computation	2.91 −.58 2.33	3.23 −.93 2.30	3.14 −.55 2.59					
Arith. Reas.	1.91 −.24 1.67	1.77 +.10 1.87	2.11 −.16 1.95	2.81 −1.88 .93				
Sci. Inf.	1.32 −.03 1.29	.47 +.98 1.45	.40 +.77 1.17	2.48 +.16 2.64	1.94 −.41 1.53			
Hist. +Lit.	1.61 −.31 1.30	.79 +.50 1.29	1.78 −1.03 .75	2.04 +.77 2.81	1.87 −.07 1.80	1.59 −.96 .63		
Lang. Usage	1.39 −.25 1.14	.97 −.06 .91	1.77 −.98 .79	2.33 +.05 2.38	1.30 +.74 2.04	.56 +1.12 1.68	.87 +.56 1.43	
Spelling	1.32 +.53 1.85	1.86 −.25 1.61	1.69 − .26 1.43	3.05 −.82 2.23	2.27 +.06 2.23	.71 +1.83 2.54	1.36 +1.01 2.37	2.11 −.78 1.33

The three entries in each cell in Table II are in order: (1) The nature factor a, constant throughout life, contributing to adult idiosyncrasy; (2) the nurture factor u, the sum of all nurture factors preceding adulthood, contributing to adult idiosyncrasy; and (3) the sum of the two $(a + u)$, which is the total adult idiosyncrasy measure. The values of a, u, and $(a + u)$ will vary together, depending upon the reliability of the test employed. These values may, however, be compared with each other, for throughout, the situation has been reduced to one involving equal reliabilities. Thus, we find in the first cell $(a + u) = .92$ and in the second $(a + u) = .79$. Though we cannot readily interpret the absolute values .92 and .79, we may compare the two values and conclude that in sense differences there is slightly greater adult idiosyncrasy between the Paragraph Meaning and the Sentence Meaning functions than between the Paragraph Meaning and the Word Meaning functions.

If we average the eight different a values found for a single function we will have a measure of the innate distinction of that function from all the other functions measured. This has been done and the results recorded in Table III, first, however, weighting the Paragraph Meaning, Sentence Meaning, and Word Meaning values one-third each, as these three functions are known to be intrinsically very similar to each other.

TABLE III

AVERAGE INNATE IDIOSYNCRASY AND AVERAGE ADULT IDIOSYNCRASY VALUES

	P. M.	*S. M.*	*W. M.*	*Comp.*	*A.R.*	*S. I.*	*H+L.*	*L. U.*	*Spell*
Average a	1.68	1.49	1.75	2.63	2.02	1.34	1.52	1.42	1.85
Average $(a + u)$	1.52	1.48	1.36	2.23	1.74	1.72	1.69	1.63	2.07

We gather from Table III that innately Computation is the most independent of the nine functions, but that this innate in-

dependence is considerably weakened by nurture, the average innate value 2.63 becoming the average adult value 2.23. Even so, the function remains the most independent throughout life as 2.23 is larger than the next highest $(a + u)$ value, 2.07, which is that for Spelling. It would seem that elementary school children relatively weak in Computation are especially spurred on to eliminate this weakness and bring their Computation ability up to their average level and that school children who are superior in Computation are neglected with reference to it, or especially spurred on in other subjects, thus tending to bring the function down to their average level.

Section 7.—The Educational Import of Certain Conditions Found

We may pause to ask if this leveling process in the matter of Computation is socially valuable. Questions of a common culture as they bear upon the non-vocational duties of citizenship are but little involved, while questions of vocational efficiency are intimately connected with Computation ability. History and Literature Information, in which nurture has on the whole added to innate idiosyncrasy* (average $a = 1.52$ and average $[a + u] = 1.69$) is probably more intimately connected with the universal duties of citizenship than any other school subject except, perhaps, reading. Should not our national educational institutions tend more strongly to a common level of achievement in History and Literature Information than towards such a level in Computation ability? The meeting ground of alien peoples is not in a common arithmetical ability, but in a common language (Reading, Language Usage, Spelling) and a common literary and historical knowledge and aspiration (History and Literature Information). Though the individualist

* This statement, based upon elementary school children, would probably be strengthened if high school data were available, as elimination from school during the high school period and differentiation in courses of study as regards History and Literature would increase the tendency here noted.

will object upon all occasions to a weakening of individuality by eliminating innate idiosyncrasies, and though the common culture advocate will argue for such elimination, still both should agree that from the standpoint of social structure and stability it makes a great deal of difference what the mental function is that is being considered. Apparently, judging by the important elementary school subjects, schoolmen resent oddity wherever it appears, as would a street mob or a class-conscious gang. Due to actual innate idiosyncrasy, to objectivity of measurement, and to time devoted to the subject, oddity in Computation ability is more quickly spotted than any other peculiarity; so the teacher whose Procrustean creed is "equalize, equalize" proceeds with vigor and effectiveness to the forcing-stunting task before him. The trouble is with the creed. Why equalize unless some social good result? If schoolmen confined their efforts towards securing uniformity to Language and History they would be much more effective in these fields and less restrictive to individuality. In particular, there seems to be no argument in support of the indicated lack of concern with Language Usage. Innately, individuals show little idiosyncrasy in this matter (average $a = 1.42$) but such as they do show is augmented with growth (average $[a + u] = 1.63$). To attribute this to influences outside of the school only makes clear the inadequacy of language work within the school. When the schoolmen of America seriously wish to be effective in this subject, they will find a way.

Having found fault with American school practice for failure to eliminate unevenness of development in Language Usage, it is but fair to point out that in the important socializing agency, Reading, nurture has slightly tended to eliminate original idiosyncrasies, for the average adult is a little less unevenly developed in Reading as judged by his other talents than he was innately. For the purpose of creating literary geniuses this may be looked upon as unfortunate, but for the very important purpose of creating a well-knit social structure it may be con-

sidered a happy outcome. We find in the ranks of geniuses Anatole France, Vicente Blasco-Ibanez, Robert G. Ingersoll, etc., etc. and in mediocrity the writers for and the readers of the grossly popular magazines and the Sunday supplements. An educational philosophy and practice which nourishes the former does not encourage the latter. As conscious directors of education we should recognize this and seek a policy preservative of State and not greatly destructive of originality, for a State preserved at such a cost may scarce be worth preserving.

The writer would propose a policy which preserves and utilizes individual peculiarity, except where it is established that social stability demands otherwise. The evidence of Table II points to a practice which tends to level individual talents to an individual average irrespective of remote consequences. The churchman's view of the Middle Ages, "Oddity is an evil and must be cut off," is the schoolman's view of today. Its persistence throughout the ages points to its deep root in human nature — that is, in the psychology of the one who has control of others. Not until the teacher gives up the desire to cast all others in the high mold represented by himself will the resentment of oddity cease. Is it too much to ask that the credo of the teacher include "I shall respect and endeavor to utilize to a social outcome idiosyncrasy wherever found"?

Let us examine Table II in more detail. The largest nurture factor in the table is negative; that is, it has lessened, not augmented, an original idiosyncrasy. It is the value -1.88 found in the Computation-Arithmetic Reasoning cell. Natively these two functions are wide apart ($a = 2.81$), but nurture brings them close together ($a + u = .93$). As present courses of study are constructed, this almost seems to be a necessary situation, for Arithmetic Reasoning at present seldom finds expression in the upper elementary school grades except through the vehicle of Computation. However, in the matter of guidance one must not conclude that great Computation ability is a prerequisite to high mathematical attainment in advanced fields. Computa-

tion largely drops out in the university. This probably explains the low correlation commonly found between standing in elementary and advanced mathematics. It is very possible that many a youngster has been discouraged from the pursuit of a calling involving mathematics because of his weakness in Computation, either self-recognized or pointed out by a teacher. The competent adviser is the one who sees the innate alienation between Computation and Arithmetic Reasoning, and classes as immaterial data upon the former when considering vocations demanding the latter. The student of group theory need never compete with an adding machine, and vice versa the cashier need never determine "How old is Ann?" or perform any similar reasoning feat. These two functions, married in the elementary school, become estranged and, were it not for their child — practical application — would be quite divorced in later life. As it is, neither quite respects the lineage of the other. Computation, born of routine and memory, vies with Arithmetic Reasoning, the offspring of deduction and analysis. There is a wide gamut of vocations to choose from so that an individual markedly superior in either trait, whatever his achievement in the other, should find a calling to his liking.

An interesting situation is indicated by the figures in the Paragraph Meaning-Spelling cell of Table II: $a = 1.32$; $u = .53$ and $(a + u) = 1.85$. The Paragraph Meaning is the most thought-provoking of the reading tests and therefore demands an attentive process that is not directed toward the mechanical phases of reading. Spelling ability, on the other hand, develops as a result of attention to detail. Since the nurture factor here is positive and fairly large, it seems sound to conclude that many children develop one of two attitudes toward printed matter; (a) an interest in its meaning, or (b) an interest in its structure. This development of interest in structure is further indicated by the nurture factor, $u = -.78$, found in the Spelling-Language Usage cell. There has been classification of individuals into (1) idea thinkers and (2) thing thinkers. We see that beside the

euphonic excellence of this classification there seems to be a slight warrant for it in fact. However, it is very interesting to note that the warrant is rooted in nurture and not in nature. Paragraph Meaning and Spelling are originally, as judged by the other functions studied, quite similar, for $a = 1.32$.

The developed interest in the memorizing of detail is manifest in other school subjects than spelling. When the writer was a teacher of geometry, it was frequently forced upon his attention that a certain small percentage of his classes utilized a very different set of talents than the majority. This small group attended to detail and memorized theorems and proofs and were continually perplexed when called upon for meanings. They were, in fact, decidedly annoyed that such ridiculous demands should be made upon them after a "perfect" demonstration had been given. Is it not probable that in their cases it had been discovered early in school life that memory work led to "good" recitations and the satisfactions resulting therefrom, and that in consequence a more or less generalized reaction to all school situations by this type of mental process was built up? Memory is a necessary talent for success in much of school work and it may be a sufficient talent in most of it, — even in geometry, for what teacher will fail a pupil regularly contributing "perfect" demonstrations of all set theorems, — but it is very unfortunate if it be sufficient for all scholastic demands.

The writer has also had quite a number of Chinese students in his classes and has found a goodly proportion of them brutally attacking delicate problems in analysis with a wooden memoriter tool. The most pitiful consequence of this outrage is the inability of such students to realize the enormity of their crime or the inadequacy of their method. Do we not here have a survival of the older Chinese education which "solved" all the problems of life and the secrets of God by memory? Apparently our educational practice also tends this way. The evil, — for it is an evil when a memory bludgeoning of a problem benumbs one into contentment, thereby forestalling an analytical

approach under which it would unfold, grow beautiful, and make friendly paths in the dark forests of ignorance, — has been felt by many real teachers who have cherished initiative and independence in their charges. We apparently still have much to accomplish. We may well add the following plank to our teacher's credo: "As every problem has a richer content if solved through reasoning than through memory, I shall not be content with solutions involving the lesser vision when the greater is possible."

The picture of mental structure and of social pressures contained in Table II has scarcely been outlined, but for fear that we will lose sight of the main picture if too great detail is before us only certain outstanding features of the table are mentioned here. In brief, we see an original nature quite differently endowed in lingual and numerical talents (Computation with Paragraph Meaning, Sentence Meaning, Word Meaning, Spelling give a values of 2.91, 3.23, 3.14 and 3.05 respectively); second though much less obvious or pronounced, native differences exist in reasoning, memory, and interest, the last suggested by the status of Science Information and History and Literature Information. We also see a nurture augmentation of certain differences which is probably due to direction of interest (Paragraph Meaning-Spelling, u = .53; Science Information-Spelling, u = 1.83; History and Literature Information-Spelling, u = 1.01; Sentence Meaning-Science Information, u = .98; Word Meaning-Science Information, u = .77; and Computation-History and Literature Information, u = .77). We see a partially successful elimination of Language Usage and other verbal differences (Paragraph Meaning-Language Usage, u = –.25; Sentence Meaning-Language Usage, u = –.06; Word Meaning-Language Usage, u = –.98; Spelling-Language Usage, u = –.78). We also see a pronounced lack of establishment of parallelism in Language Usage ability and non-verbal abilities (Computation-Language Usage, u = .05. This is not large, but considering that here a = 2.33, it indicates pronounced

lack of adult parallelism. Arithmetic Reasoning-Language Usage, $u = .74$; Science Information-Language Usage, $u = 1.12$; History and Literature Information-Language Usage, $u = .56$). We may thus anticipate that adults with non-verbal talents will very commonly find difficulty in properly expressing themselves.

Section 8.—A Reaffirmation of Certain Critical Points

A careful search by the writer failed to reveal definitions of the terms "nature," "nurture," "maturity," "natural unit of measurement," in connection with mental traits, of such definiteness as to be serviceable for numerical treatment. Because of this he has felt at liberty to himself define these terms quantitatively. His definitions are here repeated, for no conclusions drawn should be interpreted in the light of other definitions. It is anticipated that some will object to the definitions, but be that as it may, they must be accepted for the purpose of understanding this study.

The natural unit of measurement, or unit of difference, in a mental trait is one proportionate to a sensed appreciation of the trait, — the sensing to be done by those competent to judge the trait. The terms "nature," "nurture," and "maturity" are each defined in terms of such units.

The maturity measure for an individual is the mean measure for those of his age for the trait in question, — thus, maturity and chronological age are perfectly correlated. A subtraction of the mean score for the age, from the individual's gross score, leaves a deviation score which is independent of maturity. Nurture and nature measures are defined in connection with measures of difference between two abilities. These measures of difference show zero correlation with age and are thus independent of maturity, but the square measures of difference, or the absolute values of differences, are not uncorrelated with age. They may be divided into two portions, (a) a portion independent of age which is called the nature measure and (b) a portion

which is perfectly correlated with exposure to nurture which is called the nurture measure. The length of time and the amount that nurture affects a given mental function is obviously correlated with age.

The reader should note that a change in gross score perfectly correlated with age is called maturity and a change in difference between scores (either positive or negative) perfectly correlated with age and grade (as all the subjects of a given age were taken from the same school grade) is called nurture.. This distinction is equivalent to the assumption that maturity always works in a single direction; namely, to increase gross scores with increase in age, while nurture may work either way, and differently with different individuals, always,* however, increasing its influence with an increase in the length of time through which it acts.

* From the standpoint of social product the two following situations are equivalent, since, in each, in the group (of two) entire the native idiosyncrasy (e) is the same as the adult idiosyncrasy (f).

		Δ *Differences Between Two Functions as Given by Original Nature*	Δ *Adult Differences Due to Original Nature plus Nurture*	*Group Measure of Nurture Influence:* (*f*) — (*e*)	*Individual Measure of Nurture Influence:* (*b*) — (*a*)
1st Situation	Individual A	1.0 (a)	1.0 (b)		0
	Individual B	−1.0 (a)	−1.0 (b)		0
	Group Measure: $\sqrt{(\Delta^2_A + \Delta^2_B)/2}$	1.0 (e)	1.0 (f)	0	
2nd Situation	Individual α	1.0 (a)	−1.0 (b)		−2.0
	Individual β	−1.0 (a)	1.0 (b)		2.0
	Group Measure: $\sqrt{(\Delta^2_\alpha + \Delta^2_\beta)/2}$	1.0 (e)	1.0 (f)	0	

These definitions are thought to sufficiently correspond with an average of widely different opinions as to the meanings of the terms, to be serviceable. Accepting them, there is still ample room for certain doubts and queries. To mention just a few: (a) Is there a more meaningful basis of mental measurement than that based on the sensed difference? (b) If not, what is the best way to determine this unit? What difference in the findings would have resulted (c) had the zero point been more reliably determined and had different nurture proportions been found; (d) had different nurture values been used for the different mental functions instead of a single value for a given age for all the functions; (e) had probable errors been given throughout; etc.?

These questions cannot be answered at the present time, but the problem merits a repetition and a further investigation of all these issues. No single issue is more worthy of further research than the outstanding finding that, on the whole, nurture tends to eliminate idiosyncrasy; that, on the whole, children are less *tabulae rasae*, less products of a common mold, than adults.

Group measures of the general sort given in the next to the last column and not of the sort given in the last column are the ones determined in this study and the conclusions drawn are intended to be "social import" conclusions, and not those having "individual import." A study is under way the object of which is to determine individual measures of nurture. These will undoubtedly average somewhat larger than the group nurture measures here given.

Part II. THE IDIOSYNCRASIES OF GIFTED CHILDREN

The conclusions of Part I apply to as nearly normal groups as could well be selected. In addition, the data are available for a comparison of gifted and normal children*. Four groups

TABLE IV

Comparison of Gifted and Normal Populations

(a) 108 Gifted 7 and 8-year-olds. Mean age 8.2. Mean Stanford Binet I. Q.=151.
(b) 125 Gifted $10\frac{1}{2}$ to $11\frac{1}{2}$-year-olds. Mean age 11.0. Mean Stanford Binet I. Q. = 149.
(c) 825 Normal 8-year-olds. Mean age 8.5. Probable mean Stanford Binet I. Q. = 100.
(d) 887 Normal 11-year-olds. Mean age 11.5. Probable mean Stanford Binet I. Q. = 100.

		Total	*Para. Mean.*	*Sent. Mean.*	*Word Mean.*	*Comp-utat.*	*Arith. Reas.*	*Spell-ing*
Means	Gifted 8.2-year-olds	35	61	34	38	74	52	91
	Normal 8.5-year-olds	12	15	9	7	33	18	35
Standard Deviations	Gifted 8.2-year-olds	8.3	19	14	14	26	20	30
	Normal 8.5-year-olds	5.8	13	7	7	15	10	19
Means	Gifted 11.0-year-olds	61	93	65	68	129	98	159
	Normal 11.5-year-olds	38	58	36	39	96	59	100
Standard Deviations	Gifted 11.0-year-olds	4.9	11	10	7	16	18	24
	Normal 11.5-year-olds	7.6	16	12	11	20	17	26

* Such a comparison, following a technique somewhat different from that here used, has been made by Dr. James C. De Voss in a doctor's dissertation written at Stanford University. His conclusions are in harmony with those here reported.

will enter into the comparison. Data descriptive of these groups are given in Table IV. For brevity the groups will be referred to as "the 8-year-old gifted," "the 8-year-old normal," "the 11-year-old gifted," and "the 11-year-old normal" populations, but reference to the age means and age ranges given in Table IV will show that these are not exact designations.

The two gifted groups are very superior to normal children in their Stanford Achievement capacities. We are not here concerned with this gross difference in ability, but are interested in unevenness in ability. Measures of idiosyncrasy, i^2_{12}, have been calculated for these gifted children. For each test the mean score for normal children was subtracted from the individual's gross score to eliminate maturity, for from our definition the

TABLE V

MEASURES OF GIFTED AND NORMAL IDIOSYNCRASIES, $-\,i^2_{12}$

In each cell the value for gifted 8-year-olds is recorded in the upper left-hand corner.
The value for gifted 11-year-olds is recorded in the upper right-hand corner.
The value for the normal 8-year-olds is recorded in the lower left-hand corner.
The value for the normal 11-year-olds is recorded in the lower right-hand corner.

	Paragraph Meaning	*Sentence Meaning*	*Word Meaning*	*Computation*	*Arith. Reasoning*
Sent. Mean.	1.20 .57 1.12 1.18				
Word Mean.	.95 .57 .78 1.17	.69 .67 1.08 .68			
Computation	6.30 1.82 2.77 2.50	6.16 2.53 3.00 2.60	6.46 2.09 2.99 2.79		
Arith. Reas.	4.33 1.54 1.85 1.72	3.93 2.31 1.78 1.89	5.42 2.17 2.03 2.09	6.53 .65 2.28 1.66	
Spelling	3.64 1.49 1.33 2.02	3.74 1.76 1.80 1.70	3.29 1.17 1.62 1.53	7.07 1.76 2.80 2.58	5.73 1.96 2.29 2.28

maturity of a gifted child of a certain age is the same as that of a normal child of the same age. Thus, using not the mean of the gifted group but the mean of a normal group of the same age, product moment functions, standard deviations, etc. were calculated and the results reduced to comparable bases as before,—that is, the standard error of measurement in sense differences was made the same for all groups dealt with. Table V gives these results for the gifted group together with the comparable figures taken from Table I for the normal children.

Before examining into the general situation revealed let us consider one phase of the table wherefrom conclusions can be drawn which do not involve the major assumptions as to the equality of sense units necessary in all of the previous interpretations. If in Table IV we compare the gifted 8-year-old mean scores with the normal 11-year-old mean scores, we have values which are very nearly equal, as follows:

Gifted 8-year-olds	35	61	34	38	74	52	91
Normal 11-year-olds	38	58	36	39	96	59	100

In other words, the part of the test used in testing the gifted 8-year-olds was substantially the same part as that used in testing the normal 11-year-olds. As a consequence there is here no hazard in the assumption that standard errors of estimate are equal. The situation is similar to one in which young heavy children and older average children are weighed on the same scales, the average weights for the two groups being about the same. In fact, for the comparison of the gifted 8-year-olds and the normal 11-year-olds it would not have been necessary to express the scores in sense difference units as the original units would be the same for the two groups. Of course, expressing each in sense difference units has not changed this equality. Accordingly, when we find, as we do in the Paragraph Meaning-Sentence Meaning cell, that i^2_{12} for the gifted 8-year-olds, 1.20, is almost exactly equal to that for the normal 11-year-olds, 1.18,

we can conclude with assurance that the two groups are homogeneous with reference to unevenness in these two traits. Similar conclusions hold with reference to Paragraph Meaning-Word Meaning and Sentence Meaning-Word Meaning, but do not at all hold with reference to any other pair of traits chosen.

If we compare the upper left and the lower right-hand corner entries throughout, we will, except as between the three reading tests, find great differences in the values. Consider Paragraph Meaning and Computation. If we were to examine each gifted 8-year-old child we would find many of them very superior in Paragraph Meaning to Computation and many the reverse, while similar examination of the normal 11-year-olds would show that they tend toward much greater uniformity in these two abilities. The magnitude of this difference is peculiarly striking and undoubtedly reflects radical differences in both the nature and the nurture of the two groups.

Even more striking is the difference found in the Computation-Arithmetic Reasoning cell. Here i^2_{12} for the gifted 8-year-olds is 6.53 and for the normal 11-year-olds is 1.66. The antecedents of the gifted may be characterized as exceptional heredity, informal, varied and largely self-initiated home training, and about two years of regular schooling; while for the normal group we may expect average heredity, somewhat less varied home training, and about five years of regular schooling. In the first instance, original nature has found wide expression, and very superior arithmetic reasoning ability coupled with inferior computation ability or the reverse has blossomed like a lily by the roadside, oblivious to the fact that in the homeland of America it is only the dandelion that is expected there. As soon as this *faux pas* is discovered, teacher and pupil combine to correct it, for in another three years we find that Arithmetic Reasoning and Computation abilities have become very evenly developed, $i^2_{12} = .65$. This numerical statement is very likely an overstatement of the leveling process, for the 11-year-old gifted individuals are probably not a sampling of the same sort

as the 8-year-old gifted children. It is very possible that in part this difference between the 8-year-old idiosyncrasy, 6.53, and the 11-year-old value, .65, is to be attributed to differences in selection and not to nurture. The very great difference here found and also found in other cells between the gifted 8 and the gifted 11-year-old i^2_{12} values suggests that an older child gifted in some one line and not in other lines is considered odd rather than gifted, and therefore not found in the selection which was in part based upon teachers' judgments, while the 8-year-old highly gifted in any line is selected even though relatively inferior in some second line.* It does not seem reasonable to attribute the major portion of the differences found to this cause, so that apparently the instruction of gifted children from age 8 to 11 has been very effective in eliminating idiosyncrasy. For the gifted children these three years are very much more years of public school influence than the earlier years, so that here with the gifted, as earlier with the normal, we find evidence of the public school very effectively eliminating oddity.

The idiosyncrasy values for the gifted 11-year-olds average smaller than for the normal 11-year-olds. Thus in the matter of idiosyncrasy these gifted children are more like normal 14-year-old children, just as they are in the matter of gross score, than they are like 11-year-olds. Whatever argument there is for developing a dead level of ability with mediocrity, it is surely less forceful with reference to gifted children, all of whose abilities are well above the average. These gifted children are not undifferentiated paragons of ability because they were born that way, nor because they were that way at age 8, but because something has happened between ages 8 and 11. Probably the thing that has happened is that they have succumbed to the demands and blandishments of a school system and a school influence reacting on the home, organized for the mediocre

* Dr. Terman, who is most familiar with the means employed in selecting the gifted children, thinks that I have overstated the importance of differences in the selective tendencies affecting the younger and the older age groups.

and administered by plainsmen — lovers of the level. A hardy mountaineer, accustomed to pin his faith on a single staff, to trust himself in traveling uncharted routes, is the proper guide for those who we hope will follow in his footsteps to the end of the path and then blaze a trail beyond.

When established social routines are forsworn we find chess-playing and musical geniuses of ages 8 or 10; and when we do, we immediately pity the poor distorted creatures and exercise our beneficent influence to round them out, and we succeed so well that these youthful geniuses are seldom heard of in later life. Only a few creep through the barrage, and a few others avoid it by being neglected waifs as children and shunned as peculiar as adolescents, so that it is only when full blown that they are "discovered" as saviors of mankind. Why should we not have hundreds of such where we now have tens? We, the schoolmen of America, can have, if we open our eyes to the rare growth about us and if we compose our hearts to the concept that frequently in the small frames that pass in review before us are greater minds and larger visions than our own. The writer, for reasons unknown to the teacher, visited an eighth grade class in which there was a 190 I. Q. youngster. Here, by a generous estimate, was a 100 I. Q. teacher bullying all, including this particular child who quietly and without provoking a responsive chord, gave expression to mental processes as straightforward, ingenious, and brilliant as a tender blade of grass with the morning dew upon it. Let us not prejudge the teacher, but we may doubt if his nightly prayer concluded, "— and, Heavenly Father, I thank thee for permitting me this day to enjoy the fragrance of a great soul."

Only we as teachers have this privilege, and, at our option, we may enjoy and profit by the attempt to understand the noble cast of character entrusted to our keeping, or we may search and discover trifling shortcomings — a weakness here in computation, there a word misspelled, — and devote our high talents toward correcting the offense.

Part III

APPENDIX A

Variabilities of Groups in Sensed Differences and the Relation Between the Units Used

It is a simple matter to express an individual deviation score as a multiple of the standard deviation of the group. Thus, if we can but ascertain the relation between the variabilities of our three groups in terms of sensed differences we can express all deviation scores in terms of these differences or, at least, in terms proportional to them, which is just as satisfactory. The reaction of teachers of different grades to their pupils in the matter of promotion makes possible a comparison of the relative spread in sensed differences of the different grade groups.

We will argue that a pupil in a given grade who deviates some definite number of sensed differences from the mean of the group is, because of this fact, relegated to a different group; that he is either not promoted, or demoted, or doubly promoted, etc. Thus, whether a third or an eighth grade pupil be so a-typical as not to be regularly promoted, we will consider that he is removed from the mean of the group by an amount which is in excess of some definite number of sensed differences, the number being the same for each grade.

Let us assume the word "exceptional" to mean the same number of sense differences, whether 8, or 14-year-olds are dealt with, and assume a similar constancy in the meaning of the term "very exceptional." If an 8-year-old must be "very exceptional" before he is not regularly promoted, it is here assumed that a 14-year-old must also be "very exceptional" before he is not regularly promoted. If an 8-year-old needs to be "very

exceptional," and a 14-year-old only "exceptional" for each to not be regularly promoted, then the units here employed are not proportional to true sense differences. The writer, however, sees no adequate reason for thinking that this last-mentioned situation holds, and he believes that the units employed are substantially proportionate to sense differences. Two supplementary investigations mentioned in the footnote on page 42 tend to confirm him in this view.

All the available promotion records for a number of years for San Jose, California, 8.50–9.00-year-olds were investigated with the result shown in Table A.

TABLE A

Grade in which found at age 8.50–9.00	low 1	high 1	low 2	high 2	low 3	high 3	low 4	high 4	low 5	
No. found	16	56	101	179	204	140	44	10	1	751
Per cent of 8-yr.-olds not regularly promoted	12	11	11	10	7	8	14	10	0	

Smoothed value for grade 2.8 is taken as 8.4%

If the distribution of general scholastic ability of 8.50–9.00-year-olds in grade 2.8 is taken as normal, then reference to a table of the normal probability integral shows that it is necessary to go 1.728 standard deviations up and down from the mean to leave at the extremes a total of 8.4 per cent. This distance, 1.728 standard deviations, is such a sense difference, let us call it S, that when pupils are recognized as lying farther than this above or below the mean they are thrown out of the regular line of promotion.

A similar procedure for the 11-year-olds yields the data of Table B.

TABLE B

Grades in which found at age 11.50–12.00	low 1	high 1	low 2	high 2	low 3	high 3	low 4	high 4	low 5	high 5	low 6	high 6	low 7	high 7	low 8	
No. found	1	14	17	47	70	87	101	138	201	220	212	150	63	25	3	1349
Percent of 11-yr.-olds not regularly promoted	0	0	6	6	6	11	12	8	8	5	6	5	6	4	0	

Smoothed value for grade 5.65 is taken as 6.4%

TABLE C

Grades in which found at age 14.50–15.00	high 3	low 4	high 4	low 5	high 5	low 6	high 6	low 7	high 7	low 8	high 8	low 9	high 9	
No. found	2	6	9	13	26	41	61	103	127	191	58	no data	no data	637
Per cent of 14-yr.-olds not regularly promoted	0	50	0	15	0	17	7	9	6	3	7			

Smoothed value for grade 8.55 is taken as 5.2%

Beyond 1.852 standard deviations from the mean yields 6.4 per cent. Accordingly 1.852 of these 11-year-old standard deviations is also S. Finally, for the 14-year-olds we have Table C (see p. 40). Beyond 1.943 standard deviations from the mean are found 5.2 per cent of the population, so that 1.943 of these 14-year-old standard deviations is also S. Using notation of the type "$x \equiv y$" to mean that x and y are equivalent scores when each is interpreted in terms of sensed differences, we have:

$$S \equiv 1.728 \text{ eight-year-old standard deviations}$$
$$S \equiv 1.852 \text{ eleven-year-old standard deviations}$$
$$S \equiv 1.943 \text{ fourteen-year-old standard deviations}$$

In this statement "8-year-old standard deviations" refers of course to the standard deviation in sensed differences of general scholastic ability or achievement of 8-year-olds found in a particular grade location, — grade 2.80 or thereabouts, — and not to the standard deviation of a complete or random selection of 8-year-olds. Similarly for the 11- and 14-year-old standard deviations. If σ_1, $\boldsymbol{\sigma}_1$, and Σ_1 are standard deviations in terms of sensed differences we may express the same relation by

$$1.728\sigma_1 = 1.852\boldsymbol{\sigma}_1 = 1.943\Sigma_1. \qquad [14]$$

Let us assume that the teachers of the pupils in the three groups are equally excellent judges and that the reliability of their judgments for the 14-year-olds in the eighth grade is .70.* We may then estimate the true standard deviation, Σ_∞, of the

* This reliability is a little higher than has usually been found for a single grade group, but it is thought that age is consciously considered by certain teachers and not by others, resulting sometimes in the promotion of dull, old pupils and the retention of young, bright pupils. Such a procedure would lower an obtained measure of reliability. As all of our 14-year-olds are at grade for their age this factor is not here present, so that it seems reasonable to expect that the reliability of teachers' judgments would be somewhat higher than is commonly found for a typical grade group.

14-year-old age group in sense differences by the formula $\Sigma_\infty = \Sigma_1 \sqrt{R_{1I}}$. * This yields:

$$\Sigma_\infty = \Sigma_1 \sqrt{.70} = .837\Sigma_1.$$

Doing the same for the other age groups we obtain:

$$\boldsymbol{\sigma}_\infty = \boldsymbol{\sigma}_1 \sqrt{.7274}\ \dagger = 1.049\Sigma_1 \sqrt{.7274} = .895\Sigma_1$$
$$\sigma_\infty = \sigma_1 \sqrt{.7628}\ \dagger = 1.1245\Sigma_1 \sqrt{.7628} = .982\Sigma_1$$

These equations provide us with the ratios‡:

$$\sigma_\infty : \boldsymbol{\sigma}_\infty : \Sigma_\infty = .982 : .895 : .837 \qquad [15]$$

From these we note that the three groups are not equally variable in true sense differences, the younger group being the most variable and the older the least.

Assuming that the 6-test Stanford Achievement total score is a good measure of such general all-round scholastic ability as teachers are concerned with in the matter of promotion we should find the same ratio between estimated true Stanford Achievement variabilities for the three groups as ratios [15],

* See Kelley, *Statistical Method*, Formula [166].

† If the reliabilities of teachers' estimates are equal in the different groups, then the standard errors of estimate are equal, giving:

$$\Sigma_{1\cdot\infty} = \boldsymbol{\sigma}_{1\cdot\infty} = \sigma_{1\cdot\infty}$$

or

$$\Sigma_1\sqrt{1-\mathrm{R}_{1I}} = \boldsymbol{\sigma}_1\sqrt{1-\mathbf{r}_{1I}} = \sigma_1\sqrt{1-r_{1I}}.$$

Knowing the relationship as given in equation [14] between Σ_1, $\boldsymbol{\sigma}_1$, and σ_1, and taking R_{1I} as equal to .70 we may determine $\mathbf{r}_{1I}$ and r_{1I}. This procedure yields the values $\mathbf{r}_{1I} = .7274$ and $r_{1I} = .7628$ here used.

‡ These ratios were investigated by two other methods, involving however somewhat questionable assumptions. The first of these assumed that the distribution of typical 11-year-olds in true achievement ability was normal and that the rate of change of the real significance of Stanford Achievement total test score units was constant in proceeding from 8 to 14-year-olds. The second involved an assumption as to the ages at entrance to school of 501,521 California elementary school children found in designated school grades by the research division of the department of Education of the University of California. The first of these methods gave an 8 to 14-year-old ratio of .947 : .837, and the second gave the ratios $\sigma_\infty : \boldsymbol{\sigma}_\infty : \Sigma_\infty =$ 1.005 : .873 : .837. Accordingly the three radically different methods yielded quite similar results.

provided Stanford Achievement units are proportional to sense difference units. If this ratio does not hold we will then multiply the Stanford Achievement units for the different age groups by the appropriate constants to make it hold, and thus reduce all units to such a basis that they are proportional to sense difference units.

The reliability of the 6-test total score for grade 3 has been found to equal .95. Taking this as a cue we will assume .94 as the reliability for the 8-year group and call it ρ_{1I}, and such reliabilities, namely $\boldsymbol{\rho}_{1I} = .929$* and $\mathrm{P}_{1I} = .919$, as yield equal standard errors of estimate for the other groups. We estimate true standard deviations for the three groups as given below. In the first of these equations σ_ω is the estimated true standard deviation and σ_t is the actual standard deviation of the 8-year group in Stanford Achievement 6-test total score units, while ρ_{1I} is the reliability of the actual total scores for this group. The symbols in the second and third equations have similar meanings with reference to 11- and 14-year-olds.

$$\sigma_\omega = \sigma_t \sqrt{\rho_{1I}} = 57.75 \sqrt{.94} = 55.99$$
$$\boldsymbol{\sigma}_\omega = \boldsymbol{\sigma}_t \sqrt{\boldsymbol{\rho}_{1I}} = 76.38 \sqrt{.929} = 73.60$$
$$\Sigma_\omega = \Sigma_t \sqrt{\mathrm{P}_{1I}} = 73.04 \sqrt{.919} = 70.03$$

Comparing these values of σ_ω with ratios [15] we immediately see that the standard deviations in terms of the Stanford Achievement test units are not proportional to the standard deviations in sense difference units. To make them so let us keep Σ_ω as it is and multiply $\boldsymbol{\sigma}_\omega$ by **k,** the requisite multiplying factor of the 11-year-old test units to make these units pro-

* We have, by equation [5],

$$\frac{\Sigma_\infty}{\sigma_\infty} = \sqrt{\frac{\mathrm{P}_{1I}}{1-\mathrm{P}_{1I}}} \Big/ \sqrt{\frac{\rho_{1I}}{1-\rho_{1I}}}$$

and as we know $\Sigma_\infty/\sigma_\infty$ and have assumed ρ_{1I} to equal .94 we may determine P_{1I}, obtaining .919. We similarly find $\boldsymbol{\rho}_{1I} = .929$.

portional to 14-year-old units when considered upon the sense difference basis. We have

$$\frac{\mathbf{k}\sigma_\omega}{\Sigma_\omega} = \frac{\sigma_\infty}{\Sigma_\infty}$$

or

$$\mathbf{k}\,\frac{73.60}{70.03} = \frac{.895}{.837}$$

that is, $\mathbf{k} = 1.017$. Similarly k is found equal to 1.468.

Thus, if we leave unchanged the total score units in the neighborhood of 56.9, which is the mean total score for the 14-year-olds, and multiply those in the neighborhood of 38.4, which is the mean for the 11-year-olds, by 1.017, and multiply those in the neighborhood of 10.9, which is the mean for the 8-year-olds, by 1.468, we will secure units throughout which are equivalent in terms of sense differences. To facilitate doing this, whatever the raw score, Tables D and E following have been built up.

TABLE D

Variability in the Neighborhood of the Total Score Given Below	*Is To Be Multiplied by the Amount Given Below To Make Units Comparable Throughout in Terms of Sensed Differences*
0	2.02
2	1.84
5	1.66
10	1.48
10.9*	1.468
15	1.33
20	1.21
25	1.12
30	1.06
35	1.02
38.4*	1.017
40	1.00
45	1.00
50	1.00
55	1.00
56.9*	1.000
60	1.00
etc.	1.00

In this table the items starred constitute the data as just determined and the other items are estimates resulting from interpolation and extrapolation in what was thought to be a reasonable manner. Because distances above zero, earlier determined* to be in the neighborhood of total score minus one, are needed to determine the environmental factors referred to as q, Table E, which is an amplification of Table D, is given.

TABLE E

Stanford Achievement 6-Test Total Raw Score		*Amounts to be added for the addition of one in the raw score to reduce to a sense difference basis throughout*	*Resulting distance above zero of total score in units proportional to sense differences throughout*
	−1		0.00
	0	2.02	2.02
	1	1.92	3.94
	2	1.84	5.78
	3	1.77	7.55
	4	1.71	9.26
	5	1.66	10.92
	6	1.62	12.54
	7	1.58	14.12
	8	1.54	15.66
	9	1.51	17.17
	10	1.48	18.65
M. of 8-year group	10.9		19.95
	11	1.45	20.10
	12	1.42	21.52
	13	1.39	22.91
	14	1.36	24.27
	15	1.33	25.60
	16	1.30	26.90
	17	1.27	28.17
	18	1.25	29.42
	19	1.23	30.65
	20	1.21	31.86
	21	1.19	33.05
	22	1.17	34.22
	23	1.15	35.37
	24	1.13	36.50
	25	1.12	37.62

* See Revised Stanford Achievement Test Manual, 1926.

TABLE E—*Continued*

Stanford Achievement 6-Test Total Raw Score	*Amounts to be added for the addition of one in the raw score to reduce to a sense difference basis throughout*	*Resulting distance above zero of total score in units proportional to sense differences throughout*
26	1.10	38.72
27	1.09	39.81
28	1.08	40.89
29	1.07	41.96
30	1.06	43.02
31	1.05	44.07
32	1.04	45.11
33	1.03	46.14
34	1.02	47.16
35	1.02	48.18
36	1.01	49.19
37	1.01	50.20
38	1.01	51.21
M. of 11-year group 38.4		51.61
39	1.00	52.22
40	1.00	53.22
.	.	.
.	.	.
.	.	.
M. of 14-year group 56.9	1.00	70.12
.	.	.
.	.	.
.	.	.
Adult M. 62.8	1.00	76.02

APPENDIX B

Basic Data for Normal Eight, Eleven, and Fourteen-Year-Olds

There are three measures in each of the cells, except the diagonal cells, in Table F on page 47, for the normal 8-year-olds. The first measure is r_{12}, the raw correlation obtained between the tests indicated in stub and caption; the second is $r_{\infty\omega}$, the estimated true correlation between the same variables; and the third measure is i^2_{12}, a measure of the average tendency of

individuals in the group to possess different degrees of ability in the two subjects when the reliability of each of the measuring devices is .85. The absolute size of the i^2_{12} measures depends upon the reliability .85, but ratios between them are independent of it. The entries in the diagonal cells are reliability coefficients for the tests and group in question. These reliability coefficients were determined by assuming that for each test the standard error of estimate was equal to that found in the high second grades of four California school systems where both forms of the Stanford Achievement Test had been given and reliabilities calculated therefrom.

Tables G and H provide similar data for normal 11-and 14-year-olds.

TABLE F

BASIC DATA FOR NORMAL EIGHT-YEAR-OLDS

	Para. Mean.	*Sent. Mean.*	*Word Mean.*	*Computation*	*Arith. Reas.*	*Spelling.*
Paragraph Meaning	.88					
Sentence Meaning	.71 .80 1.12	.88				
Word Meaning	.77 .86 .78	.72 .81 1.08	.90			
Computation	.42 .51 2.77	.39 .47 3.00	.40 .47 2.99	.78		
Arithmetic Reasoning	.56 .67 1.85	.57 .69 1.78	.53 .64 2.03	.46 .60 2.28	.77	
Spelling	.67 .77 1.33	.60 .68 1.80	.63 .71 1.62	.42 .51 2.80	.49 .60 2.29	.88

TABLE G

BASIC DATA FOR NORMAL ELEVEN-YEAR-OLDS

	Para. Mean.	*Sent. Mean.*	*Word Mean.*	*Comput.*	*Arith. Reas.*	*Sci. Inf.*	*Hist. +Lit.*	*Lang. Usage*	*Spelling*
Para. Mean.	.84								
Sent. Mean.	.59 .74 1.18	.75							
Word Mean.	.65 .74 1.17	.71 .85 .68	.92						
Computation	.37 .45 2.50	.34 .43 2.60	.34 .39 2.79	.82					
Arith. Reas.	.49 .62 1.72	.44 .59 1.89	.45 .54 2.09	.50 .64 1.66	.75				
Science Inf.	.59 .71 1.33	.59 .75 1.13	.70 .80 .92	.35 .43 2.59	.50 .64 1.66	.83			
History and Lit. Inf.	.60 .69 1.40	.61 .75 1.13	.69 .76 1.07	.37 .44 2.56	.50 .60 1.82	.68 .79 .94	.88		
Lang. Usage	.56 .73 1.22	.57 .80 .93	.60 .76 1.10	.36 .48 2.37	.44 .60 1.80	.54 .71 1.32	.56 .73 1.25	.70	
Spelling	.49 .56 2.02	.52 .63 1.70	.61 .66 1.53	.37 .43 2.58	.43 .50 2.28	.49 .57 1.95	.49 .55 2.04	.52 .65 1.58	.90

TABLE H

BASIC DATA FOR NORMAL FOURTEEN-YEAR-OLDS

	Para. Mean	*Sent. Mean.*	*Word Mean.*	*Com-put.*	*Arith. Reas.*	*Sci. Inf.*	*Hist. +Lit.*	*Lang. Usage*	*Spell-ing*
Para. Mean.	.72								
Sent. Mean.	.61 .79 .85	.83							
Word Mean.	.67 .85 .59	.74 .88 .47	.86						
Computation	.29 .40 2.39	.32 .41 2.38	.27 .34 2.62	.74					
Arith. Reas.	.43 .58 1.69	.43 .54 1.83	.42 .52 1.90	.56 .75 1.00	.76				
Science Information	.51 .66 1.34	.54 .66 1.37	.61 .72 1.11	.27 .34 2.63	.48 .61 1.56	.83			
History and Lit. Inf.	.54 .67 1.32	.60 .69 1.25	.70 .79 .82	.26 .31 2.75	.46 .55 1.80	.71 .82 .71	.91		
Lang. Usage	.50 .71 1.16	.58 .77 .92	.60 .78 .86	.29 .41 2.38	.37 .50 1.99	.46 .60 1.60	.52 .65 1.39	.69	
Spelling	.46 .60 1.59	.48 .59 1.63	.53 .64 1.45	.34 .44 2.24	.33 .42 2.33	.33 .40 2.40	.37 .43 2.29	.50 .65 1.39	.81